DOCUMENT RESUME

ED 212 999 CS 006 521

AUTHOR Raphael, Taffy E.; Wonnacott, Clydie A.
TITLE The Effect of Metacognitive Awareness Training on Question-Answering Behavior: Implementation in a Fourth Grade Developmental Reading Program.
PUB DATE Dec 81
NOTE 30p.; Paper presented at the Annual Meeting of the National Reading Conference (31st, Dallas, TX, December 2-5, 1981). For related document see CS 006 520.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Content Area Reading; Grade 4; *Inservice Teacher Education; Intermediate Grades; *Metacognition; *Program Effectiveness; *Questioning Techniques; Reading Ability; *Reading Instruction; Reading Programs; *Reading Research; Student Reaction; Teaching Methods; Test Format
IDENTIFIERS *Question Types; Reading Strategies; Text Structure

ABSTRACT

A study investigated the effects of a teacher inservice program on teaching fourth grade students the functional relationship between questions and the response information to which they refer. Ten teachers and 180 of their students participated in the study. Three of the teachers participated in a traditional half-day teacher workshop on question-answer relationships, three in a more extensive training program that included specific materials and weekly monitoring with feedback, two in practice-only classrooms, and two in no-treatment classrooms. Pretest and posttest data examined students' performance in identifying and responding to (1) text explicit questions (with response information explicitly stated in the same sentence as the question information), (2) text implicit questions (requiring integration of text information across sentences), and (3) script implicit questions (requiring the readers to rely on their knowledge base, or schemata, for response information). Results of ten weeks of instruction revealed that high ability students outperformed average and low ability students, and that performance on text based questions was higher than performance on script based questions. Students in both training groups were superior to the control group students in the quality of their responses. Students whose teachers had received more extensive and programmatic training were better able to identify the question-answering strategy required by each question, though they did not differ in their internal consistency or response quality. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED212999

The Effect of Metacognitive Awareness Training on Question-Answering Behavior: Implementation in a Fourth Grade Developmental Reading Program

Taffy E. Raphael and Clydie A. Wonnacott

University of Utah

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Taffy E. Raphael
Clydie A. Wonnacott
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Running Head: Metacognitive Awareness in Question-Answering

Paper presented at the National Reading Conference, Dallas, Texas, December, 1981. We would like to thank Theresa Cryns and Susan Moody for their assistance in preparing materials. Special thanks go to Jean McKinney for her extensive contribution to the data analyses. In addition we wish to thank Kathleen Kruetzer, Ruth Harper, Robert Porter, and their students of Draper Elementary School; Sharon Hansen, Sharon Newbold, Karen Stierns and their students of Sprucewood Elementary School; and the teachers and students of Altara Elementary School for their participation in the study. This research was supported by both the University Research Committee and the Bureau of Educational Research, University of Utah.

CS006527

The Effect of Metacognitive Awareness Training on Question Answering Behavior: Implementation in a Fourth Grade Developmental Reading Program

Two of the most important considerations of instructional researchers should be: (1) theoretical justification for the research and (2) implementation of the research findings into traditional classroom curricula. Unfortunately, many of the instructional practices observed in schools today are not supported by research (e.g., the teaching, or lack thereof, of comprehension as observed by Durkin, 1980). This complaint, common among researchers is countered by a parallel complaint among practitioners: that much of the instructional practices suggested by research have little or no utility for the classroom. The purpose of this paper is to describe an instructional study which attempts to be responsive both to theoretical concepts--metacognition and questioning--and to the practitioner.

Research in the analysis of questioning has traditionally fallen into three categories: (1) the categorization of questions into taxonomies, (2) the study of the proportion of questions asked from the various taxonomy categories, and (3) the examination of the educational value of the question as a tool to enhance memory for and comprehension of prose. This research suggests that a variety of question types exist, each type with an implied set of processing demands and strategies (Bloom, 1956; Barrett, 1976; Pearson & Johnson, 1978). Additionally, research reveals that question-answering during and following the reading of prose generally requires responses to more text-based, literal questions than to knowledge-based, inferential

questions (Guszak, 1966; Chou-Hare & Pulliam, 1980). Also, questions have been shown to consistently enhance learning from text (Rothkopf, 1967; Swenson & Kulhavy, 1974; Rickards, 1979). If a number of question types exist, each requiring a different question-answering strategy, and if questions enhance learning from text, investigating the possibility of direct instruction in strategies for answering questions becomes important.

With a paucity of literature specific to the teaching of children's question-answering strategies the possibility of training strategy use in general was considered. The literature in this area, based upon work in metacognition or knowledge of and control over one's own learning process (Brown, 1975) has been encouraging. It has indicated that children of minimal and low abilities have been successfully trained to implement strategies for enhancing memory (Brown & Barclay, 1976; Flavell & Wellman, 1977) and strategies for monitoring comprehension (Wong and Jones, Note 1). From this research, we infer that training in the selection and use of question-answering strategies may be possible. Since questions are a prevelant phenomenon and a useful tool for enhancing learning, training students in question-answering strategies becomes not only a possibility, but also an important instructional concern.

Teaching children to recognize the functional relationship between a question and the response information necessary for appropriately answering the question became easier with the introduction of a taxonomy of questions by Pearson and Johnson (1978), which no longer considered questions as isolated units. Rather, questions types were identified on

the basis of the relationship between the question, the text to which it refers, and the knowledge base of the reader. Their three category taxonomy includes text explicit, text implicit, and script implicit questions. These three catogories have been identified respectively by criteria such as whether a question has response information explicitly stated within the same sentence as those constituents used to create the question (e.g., John rode a horse. Who rode a horse? John); whether the question has response information stated in the text, but requires integration of that text information across sentences, paragraphs, or pages; or whether the question requires the readers to rely on their knowledge base for providing accurate response information.

To determine whether or not the use of question-answering strategies differentiated between successful and unsuccessful readers, Raphael, Winograd, and Pearson (1980) examined the performance of fourth, sixth, and eighth grade students in responding to text explicit, text implicit, and script implicit questions. They found that high ability students were more consistent in selecting appropriate strategies based upon thier recognition of the type of question asked. Text-search strategies were used for TE and TI questions, while knowledge base-search strategies were used for SI questions.

A Training Program

A program designed to train students to recognize the three question answer relationships (QARs) proposed by Pearson and Johnson (1978) was created by Raphael (1981) and successfully implemented with groups of fourth, sixth, and eighth grade students. The program was conducted over a period of four days, moving from brief to longer

passages, and from group to independent work. As a result of participation in the program, trained students performances were enhanced relative to a control group which received the definitions of the three types of questions with no specific training.

These two studies indicated that the use of question-answering strategies did differentiate between the skilled and less skilled readers, and that the performance of less skilled readers was particularly enhanced following instruction in the use of QAR recognition to select appropriate question-answering strategies. The next issue, considered in this study, was the applicability of the training program in the typical classrooms. Four research questions were asked: (1) What level of inservice training for teachers is necessary for successful implementation of the QAR instructional program (i.e., is program success a function of inservice training alone or inservice training with provision of instructional materials and monitoring of program implementation)? (2) Will strategy training with students transfer to situations where strategy use must occur spontaneously (i.e., if students are not cued to use the specific QAR task introduced during training)? (3) Does performance using the QAR task differ from performance that does not cue students to the task? and (4) What length of student instruction is optimal for enhancing fourth grade students question-answering performance? It was predicted that the inservice training with the materials and monitoring would be superior to inservice alone. Second, it was expected that instruction would facilitate children's performance in three areas: sensitivity to the demands of the questions, consistency between strategy selection and actual stra-

tegy use, and response quality. Further, it was hypothesized that the task of inducing QAR selection may be needed to improve students' quality of responses.

Method

Subjects

Participating in the study were ten fourth grade teachers and 180 of their students from a semi-rural western community. Six of the teachers participated in inservice training sessions, three each at two different levels of training. The first group received training using a specific set of instructional materials. They were instructed in the exact use of the materials, received weekly observations and feedback on the success of the program implementation, and were assisted by researchers in monitoring the performance levels of individual students, identifying those students who needed further assistance. The second three teachers received a typical half-day inservice session with instructions on how to prepare materials from those naturally occurring in the classroom. These teachers were asked to implement the program as they chose, and to keep a record of the kinds of instruction which they provided. The remaining four teachers formed two control groups.

The participating students were selected within schools from a population of 280 in three comparable elementary schools, one for each of the two training groups and one control school. Schools were open in structure and used a team teaching approach. Thus each condition could exist only within a single school. First, students in the two training schools who had not participated in the complete program of instruction were eliminated from the study. Then, the remainder of the students

from the two training schools and all students from the control school were ranked by ability levels on five measures: standardized reading vocabulary and comprehension scores from the ITBS achievement test, a decoding rate and error rate from a word identification list test, and teacher judgment. Fifteen students from each of three ability levels were selected at random in each of the two training group schools, forming the two experimental treatment groups. In addition, fifteen students from each of the three ability groups were selected at random from two sets of two classrooms in the control school. Thus, 45 were chosen across two classrooms to form the Practice Only control group and 45 were chosen from the other two classrooms to form the No-treatment control group. Preliminary analysis revealed no difference between the two control groups; therefore, for the final analyses, fifteen students were randomly selected from each of three ability levels across both of the control groups to form the control group used throughout all other analyses.

There were significant differences in ability between the control and the two training groups, F (2,125 = 4.63, $p < .05$, with the Control group (M = 71.74) at a higher level than both the Training 1 (M = 65.49) and the Training 2 (M = 60.57) groups. This was adjusted using analyses of covariance when comparisons across, rather than within, groups were performed.

Materials

Instructional Materials. Three student workbooks were created for use over three class periods within a single week. These materials, designed to teach student to recognize text explicit, text implicit, and

script implicit QARs (mnemonics for the children were Right There, Think and Search, and On My Own for each respective category) were identical to those used by Raphael (1981). Teachers from Training 1 were provided with transparencies to accompany the students' workbooks and detailed directions in terms of potential dialogue with and feedback to the students. In addition to the materials to be used in the introductory week, teachers were also provided with eight 250-word passages with two questions from each QAR category per passage to be used once a week with their students.

The teachers in the Training 2 group received no particular materials to be used with students but, rather, had modeled for them examples of introductory lessons. Therefore, materials used with these students varied according to the teachers' planning. The three teachers in this group taught on the same team and shared most of their plans and activities.

Testing Materials. Four passages of 600 - 800 words in length were developed on topics familiar to fourth grade students as determined in a pilot study. The passages were created using modifications from passages found in trade books and basal readers. Each passage had eighteen corresponding comprehension questions, six each from the TE, TI, and SI categories.

Teacher Questionaire. An informal teacher questionaire was developed to assess teachers' attitudes towards the overall program. The questions were divided among four areas: the extent to which the inservice training was reasonable and adequate, the extent to which the instructional program for the students was reasonable and adequate, the

effect the instruction may have in transfer to content area subjects, and the extent to which the teachers would implement the program again when it would no longer be required by participation in a research study.

Design and Procedure

A 3 X 3 X 3 mixed experimental design was created with the between-subjects factors of treatment (Training 1, Training 2, and Control) and ability (High, Average, and Low), and the within-subjects factor of QAR category (TE, TI, and SI). Students in Training 1 received their instruction over a ten week period, consisting of a week of intensive training, then one session per week of maintenance lessons. Students in Training 2 also received their instruction over a ten-week period, but did not have the systemmatic program of the first group. Students in the Practice Only control group received all the maintenance materials. Their teachers had been told that the materials themselves were of value and should be administered on a weekly basis (which corresponded to that of the Training 1 group). Finally, the students in the No-Treatment control group participated only in the testing at the time of the 10th week of instruction for the other two groups. Recall that because there were no significant differences between the two control groups, students were selected randomly from the two groups to form a single Control group.

Experimental Tasks. The experimental task to which all students responded involved reading experimental passages and responding to eighteen experimental questions, six from each of the QAR categories, for each of the passages. In addition to reading the passages and

responding to the questions, students in the training groups were instructed to identify the questions by QAR categories. The following example represents a typical question and the response task used by the training groups:

What is one way a fossil is made?

RIGHT THERE _____

THINK AND SEARCH _____

ON MY OWN _____

Trained students were directed to read the question, then determine the QAR category and implied question-answering strategy used to locate the response information, providing the response next to their QAR choice in support. Thus, students in the control conditions read passages and provided responses to their corresponding questions. Students in the two training groups read the same passages, on one providing only the responses as did the control groups, on the others identifying all questions by QAR category as well as providing the correct response.

Testing Procedures. Testing occurred at four points in the research study (see Table 1): (1) pre-training for both Training 1 and Training 2, (2) Post-Intensive Training after the first week for Training 1, (3) after 10 weeks of instruction for both Training 1 and Training 2, and (4) after 15 weeks of participation for Training 1. It was at the third point that both control groups were tested. This schedule provided five important comparisons:

1. Pretesting of the two training groups: Passage A-
2. Effect of the Intensive Training week, Training 1 vs (untrained as yet) Training 2: Passage B+ versus Passage B-

3. Effect of training in a transfer task (uncued setting): Passage D-
4. Effect of training in the cued condition: Passage C+ (Training groups) versus Passage C- (Control group)
5. Effect of cued versus uncued within a single group: Passage D- versus Passage D+ (Training 1)

Dependent Measures

Three independent measures were created based upon students' ability to identify questions by QAR category and/or by the quality of their responses.

Hits. To assess students' sensitivity to task demands, scores were created from the number of accurate identifications of the QAR category. Students received a point for each TE question identified as Right There, TI question identified as Think and Search, and SI question identified as On My Own, with a total of six correct possible for each QAR category on each passage.

Matches. To assess the students' internal consistency in strategy selection and strategy use, a matrix was created with credit given when questions identified as TE or TI actually had text-based responses (correct or incorrect) or questions identified as SI had knowledge-based responses (correct or incorrect). Again, a maximum of six correct per QAR category was possible for each passage.

Response Quality. The students' quality of responses was assessed on the basis of both the accuracy and the completeness of the response. Extensive piloting of materials determined standards for the type and amount of information necessary for a complete and accurate response.

Students' responses in this study were compared against these standard responses.

Results and Discussion

Results will be presented and discussed in terms of the four research questions raised in this study. Thus, the first results concerned the level of inservice training necessary for successful implementation of the QAR program in a fourth grade curriculum. The second concerned evidence of transfer to an uncued situation. The third concerned the value of the QAR task itself, and the fourth concerned how much student instruction is necessary before the QAR concept can be applied in their question-answering behaviors.

The alpha level set for this study was the traditional value of .05. However, for the convenience of the reader, alpha levels of less than .01 will be reported as such.

Preliminary analyses on their correct responses to questions on two passages revealed no significant differences between the control groups, $F(1,68) = .659$, $p > .05$ for passage C, and $F(1,68) = 1.33$, $p > .05$ on passage D. Therefore, all additional analyses that were conducted used a subset of the control groups such that the number of subjects per cell was equal across training and control groups.

Level of Inservice Training. The difference in the amount or level of inservice instruction and guidance was assessed on the basis of both students' performance levels and teacher self report. Student performance levels will be addressed first, followed by the teachers' questionaire responses.

Students' performance on hits, matches, and responses was assessed

on passages C and D after 10 weeks of instruction. Passage C required students to identify the QAR categories when responding to the questions. Passage D, the transfer test, was administered to students by their teachers during one of their content lessons to avoid cuing the students in any way to use the QAR task; only responses were required of the students. Analyses of covariance revealed no significant differences between the two training groups' quality of responses on either passage C or D. However, main effects were revealed for both question type, $F(2,250) = 71.95$, $p < .01$, and ability, $F(2,124) = 5.24$, $p < .01$, on passage C (passage D will be discussed in detail in the next section). Performance was higher on SI questions ($M = 4.80$) than on both TE ($M = 3.48$) and TI ($M = 3.94$) questions. Ability differences were in the predicted directions, with high ($M = 4.53$) better than average ($M = 4.15$) and low ($M = 3.65$). A significant ability X question type interaction, $F(4,250) = 5.31$, $< .01$, further explained these results, indicating that while TE and TI differentiated between ability groups as predicted, all students performed at the same level on SI (see Figure 1). On this particular passage, SI questions were inordinately easy, hence the unexpected result of higher performance on script than text questions, contrary to what was predicted based on other literature (e.g., Hansen, 1981; Raphael, 1981).

In examining students sensitivity to demands of particular questions, revealed by their ability to identify questions by QAR category (hits) on passage C (passage D could not provide this data), significant differences were found for treatment, $F(1,82) = 9.34$, $p < .01$, with Training 1 ($M = 3.70$) performing at a higher level than did

training 2 (M = 3.05), and for question type, F(2,166) = 12.66, p < .01, suggesting that students were most able to identify SI questions (M = 3.99), and that TE (M = 3.17) were easier to identify than TI (M = 2.96). The treatment effect was further explained in the significant treatment X ability interaction, F(2,82) = 4.04, p < .05. The main effect appears to be due to the differences in performance of the high and average ability students only (see Figure 2). While all students performed at a greater than chance level (chance = 1/3 of 6 or 2.00), for students of average and high ability performance was enhanced by the more systematized program.

In examining students sensitivity to their own behavior, the "match" between the selected QAR and their actual response strategy (text-search or knowledge base-search), analysis of covariance revealed no significant differences between groups. Main effects for both ability, F(2,82) = 4.14, p < .05, and question type, F(2,166) = 18.28, p < .01, were in the expected directions. Performance was higher as a function of ability levels, M high = 4.93, M average = 4.20, M low = 4.12. Performance was higher on text (M TE = 4.69, M TI = 4.71) than script questions (M = 3.84).

Teachers' responses to the questionaire were based upon a five-point scale ranging from Strongly Disagree to Strongly Agree. Due to the small size of the group responding to the questionaire, no formal analyses were performed. However, there was the suggestion of an overall pattern in their responses. When collapsing across both groups' ratings, the scores ranged from 3.75 to 4.75, the positive or "agree" end of the scale. When separated by groups, those teachers in Training

1 strongly agreed that inservice training was adequate (5.00) and reasonable (5.00) in terms of time demands, the program procedures for instructing the students were adequate (5.00) and reasonable (5.00). Also, these teachers felt that the materials provided for student instruction were adequate (5.00). In contrast, while the Training 2 agreed on the adequacy and reasonableness of both inservice (4.50) and student instructions (4.00), they were neutral (3.00) regarding the materials provided for instruction. In an open-ended question regarding improvements to make in the program, the Training 2 group again stated the need for more training materials.

In summary, the utility of the respective levels of inservice training for teachers based upon students' performance levels suggests that overall, there are no significant differences between groups when examining response quality or sensitivity to the task. However, if one considers only the students' ability to identify questions by QAR category, the systematized approach used for Training 1 appears to be more facilitative. Given teachers' self-report, there appear to be no differences between their feelings about the adequacy and reasonableness of the program and inservice training, however, Training 2 consistently indicated a need for more instructional materials.

Evidence of Transfer. The transfer effects of QAR instruction were assessed on the basis of student performance on passage D, administered by teachers as part of their content area program. Students were in no way cued to consider QARs or strategies for answering questions. Analysis of covariance revealed significant main effects for treatment, $F(1,124) = 7.62$, $p < .01$, and ability, $F(2,124) = 3.88$, $p < .05$. Both

Training 1 (M = 4.85) and Training 2 (M = 4.95) performed at a higher level than did the control group (M = 4.36); high ability (M = 4.85) and average (M = 4.89) performed at a higher level than did the low (M = 4.41) ability students. These effects were further explained by a significant treatment X ability interaction, $F(4,124) = 2.44$, $p < .05$ (see Figure 3). The effect can be attributed primarily to the higher level of performance of average and low ability students in the training groups as contrasted with the control. In fact, training appeared to enhance performance such that average and low ability students' performance levels did not differ from that of high ability students.

Utility of Inducing the QAR Task. To assess the importance of inducing students to use the QAR strategy selection task (questions followed by Right There, Think & Search, & On My Own), two comparisons were examined. Analysis of variance was performed on passage D for Training 1 at the ten-week testing point without the QAR task following each question and at the fifteen-week testing point with the QAR task. Because of the interval of time, a slight increase in scores could be attributed to memory, a large increase would more likely be due to the use of the QAR task. The ANOVA unexpectedly revealed no significant effect for test time, $F(1,42) = 1.07$, $p > .05$. Main effects were found for ability, $F(2,42) = 9.30$, $p < .01$ and question type, $F(2,84) = 4.02$, $p < .05$. Ability differences were in the expected directions, M High = 5.17, M Average = 5.13, M Low = 4.43. Performance on text questions (M TE = 4.88, M TI = 5.09) was higher than on SI questions (M = 4.77). These findings suggest that the QAR task may actually inhibit performance since there was not even a slight increase in performance on

questions from any category over time. Further support for this suggestion stems from the results of the correct responses comparison between training and control groups on passage C and D at the third testing point. Recall that passage C required trained students to use the QAR task while passage D did not. No differences between training and control occurred when the QAR selection strategy task was induced, but effects in favor of training did occur when the task was not induced ($p < .01$). Fourth grade students may not have the requisite cognitive capacity to perform the overt task and respond to the question.

Level of Student Instruction. Analysis of covariance performed on passage B, comparing training after the intensive one-week phase to an untrained group revealed no significant effect for treatment, $F(1,82) = 2.53$, $p > .05$. Main effects were found for ability, $F(2,82) = 3.87$, $p < .05$, and question type, $F(2,166) = 60.14$, $p < .01$, in the expected directions. The means for ability were: M High = 4.65, M Average = 4.12, M Low = 3.69. The means for question type were M TE = 4.71, M TI = 4.56, M SI = 3.18. Comparing this finding to those at ten weeks would suggest that the intensive training program is not sufficient, the maintenance period a necessity. However, this conclusion must be tempered by the fact that Training 1 group was cued to use the QAR task which may have depressed their performance level.

General Discussion and Implications

Four questions were raised by this study. The first concerned the amount of inservice training required for successful program implementation, with an expectation of "the more, the better." Findings in terms of children's performance indicated that inservice training alone

is sufficient to encourage students' internal consistency in both selection and use of question-answering strategies. However, to best facilitate children's ability to correctly identify questions by QAR category, the more systematized approach was necessary. In addition, teachers in Training 2 reported a need for more materials to adequately implement the program. This would suggest that while inservice training is sufficient in most cases, the provision of materials and some feedback is a beneficial luxury.

The second question concerned the existence of transfer to non-experimental situations. That is, when students were not cued to think in terms of the three QARs, would they still demonstrate enhanced performance levels? Since the test passage used in this comparison was administered by teachers during the normal curriculum, it is unlikely that students would be cued to the experimental task. Teachers did report that students asked if they could use QARs, but the teachers remained noncommittal throughout the lesson. Effects in favor of the instructional program at either level were quite dramatic. Students appeared to have internalized the three information-seeking strategies, applying them appropriately in their class work.

The third issue questioned the need for the QAR task following each question. It was predicted that by providing the experimental task, appropriate question-answering behaviors would be cued and performance levels would increase. Perhaps the most interesting finding of the study was the unexpected inhibiting effect of the task. Raphael (1981) compared training to control groups which had received definitions of the three QARs but no specific training. She found that training was

superior to control when both groups used the QAR task. Similarly, in this study when none of the groups used the QAR task and only responded to the questions, training was superior to control. However, when trained students used the task, their performance was reduced to the same levels as the control group. Further, when exposed to a passage for a second time, using the QAR tasks only on the latter test period, performance levels did not change across time, yet, one would expect them to increase on the second exposure. These findings are not unlike those found by Raphael and Wonnacott (1981) in a study of the effect of inserted questions on fourth and eighth grade students' comprehension of prose. One manipulation required students to respond to the inserted questions either by writing (overt) or by thinking of (covert) the correct response. It was predicted that the overt condition would enhance performance. The results of the eighth grade students were as predicted, but fourth grade students' performance was superior in the covert condition. One possible explanation for these unexpected findings is that the younger students were unable to cope with the extra task demands required by overt tasks (e.g., identifying QARs or writing overt responses to inserted questions). This implies that while instruction in QAR strategy selection is valuable and improves performance, the task itself is useful only during instruction and practice. This is rather fortunate since the real world of texts and questions does not list Right There, Think and Search, and On My Own after each query.

The final question concerned the length of student instruction needed for learning to apply question-answering strategies. Raphael's

(1981) students demonstrated successful learning following one week of intensive training. Yet, her familiarity with the concepts and program may have provided an advantage unavailable to the classroom teachers. Unfortunately, because of the inhibiting nature of the QAR task, the results which apply to this question must be interpreted with extreme caution. After a week of intensive training, students in this study did not differ in their level of performance from an untrained group. This could mean they had not yet learned the task, or that they were inhibited by using the QAR task. Future research using the task on one passage and corresponding questions and without the task on another passage and question set is necessary before this question can be answered.

In summary, inservice training provided background sufficient for QAR program implementation, with a preference towards inservice and materials provision. Once the QAR strategies had been learned, performance levels were enhanced most for fourth grade students when they did not have to identify QARs as they responded to the questions. Finally, participation in eight weekly sessions following a week of intensive training appeared to be beneficial, though future research may reveal a shorter instructional period requirement.

Limitations of the Study

The limitations to this study can be thought of in terms of methodological problems or theoretical shortcomings. In the former category are such factors as only fourth grade students participated, thus limiting the generalizability of this study to other populations of school students. In addition, only four passages, all familiar, were

used which limits the generalizability to other materials. In the latter category are factors such as the wide variation of TI questions which was not delineated in this study, the problem of possible task interference in determining the amount of instruction necessary for students, and the need to consider other basic variables which may influence performance such as access to the text during question-answering.

Future Directions

Belmont & Butterfield (1977) have stressed the importance of longitudinal studies, particularly when involved with instructional issues. Still to be determined are the long terms effects of the training involving: (1) the students who had been trained and (2) the participating teachers. For the students, one should examine the extent to which they both remember and use flexible question-answering strategies. For the teachers, the issue is one of whether they continue to teach future classes of the QAR strategy use. Therefore, one should examine the level of performance of their current class as contrasted to students from a naive teacher.

Further research is also necessary to determine differences in the instructional needs of upper and lower grade school students. The QAR task may be facilitating rather than inhibiting for older students. Our on-going research suggests that younger children can only monitor a text-script dichotomy, unable to understand the TE-TI distinction. If we can instruct students to better cope with the questions they must so often face, we are providing a step forward in the application and integration of metacognition and questioning.

Reference Notes

Wong, B. Y. L., & Jones, W. Increasing metacomprehension in learning-disabled and normal-achieving students through self-questioning training, unpublished manuscript, Simon Fraser University.

References

Barrett, T. C. Taxomony of reading comprehension. In R. Smith & T. C. Barrett (Eds.), Teaching reading in the middle grades. Reading, Massachusettes: Addison-Wesley, 1976.

Belmont, J. M., & Butterfield, E. C. The instructional approach to developmental cognitive research. In R. V. Kail, Jr., & J. W. Hagen (Eds.), Perspectives on the development of memory and cognition. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Bloom, B. B., A taxonomy of educational objectives. New York: McKay Company, 1956.

Brown, A. L. The development of memory: Knowing, knowing about knowing, knowing how to know. In H. W. Reese (Ed.), Advances in child development and behavior (Vol. 10). New York: Academic Press, 1975.

Brown, A. L., & Barclay, C. R. The effects of training specific mnemonics on the metamnemonic efficiency of retarded children. Child Development, 1976, 47, 71-80.

Chou-Hare, V., & Pulliam, C. A. Teacher questioning: A verfication and an extension. Journal of Reading Behavior, 1979, 12, 69-72.

Durkin, D. What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 1978-79, 14, 481-533.

Guszak, F. J. Relationship between teacher practice and knowledge of reading theory in selected classes (Project Rep. No. 5-437). Washington, D. C.: United States Department of Health, Education, and Welfare, 1966. (ERIC ED0010191)

Flavell, J. H., & Wellman, H. M. Metamemory. In R. V. Kail & J. W. Hagen (Eds.), Perspectives on the development of memory and cognition. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

Pearson, P. D., & Johnson, D. D. Teaching reading comprehension. New York: Holt, Rinehart, & Winston, 1978.

Raphael, T. E. The effect of metacognitive awareness training on students' question answering behavior. Unpublished doctoral dissertation, University of Illinois, 1981.

Raphael, T. E., Winograd, P., & Pearson, P. D. Strategies children use when answering questions. In M. L. Kamil & A. J. Moe (Eds.), Perspectives on reading research and instruction. Washington, D. C.: National Reading Conference, 1980.

Raphael, T. E., & Wonnacott, C. A. The effect of response and type of posttest on understanding of and memory for text. Paper presented at the National Reading Conference, Dallas, Texas, 1981.

Rickards, J. P. Adjunct postquestions in text: A critical review of methods and processes. Reveiw of Educational Research, 1979, 49, 181-196.

Rothkopf, E. Z. Learning from written instructive materials: An exploration of the control of inspection behavior by test-like events. American Educational Research Journal, 1966, 3, 241-249.

Swenson, J., & Kulhavy, R. W. Adjunct questions and the comprehension of prose by children. Journal of Educational Psychology, 1974, 66, 212-215.

Table 1

Testing Sequence with Passages Used

	Testing Periods			
	Pre-Training	1-Week	10-Weeks	15-Weeks
Group				
Training 1: Inservice & Materials	A-*	B+	C+ & D-	D+
Training 2: Inservice Alone	A- & B-	--	C+ & D-	--
Control 1: Materials Only	--	--	C- & D-	--
Control 2: Untrained	--	--	C- & D-	--

* Letters indicate passage: A = Dog, B = Clown, C = Bicycle, D = Dinosaur
"+" indicates questions with use of QAR Selection Task (cf. p. X)
"-" indicates questions without use of the QAR Selection Task

List of Figures

Figure 1. Graph of the Ability X Question Type Interaction, Correct Responses, Passage C.

Figure 2. Graph of the Treatment X Ability Interaction, Hits, Passage C.

Figure 3. Graph of the Treatment X Ability Interaction, Correct Responses, Passage D.

FIGURE 1. ABILITY X QUESTION TYPE INTERACTION, CORRECT RESPONSES, PASSAGE C.

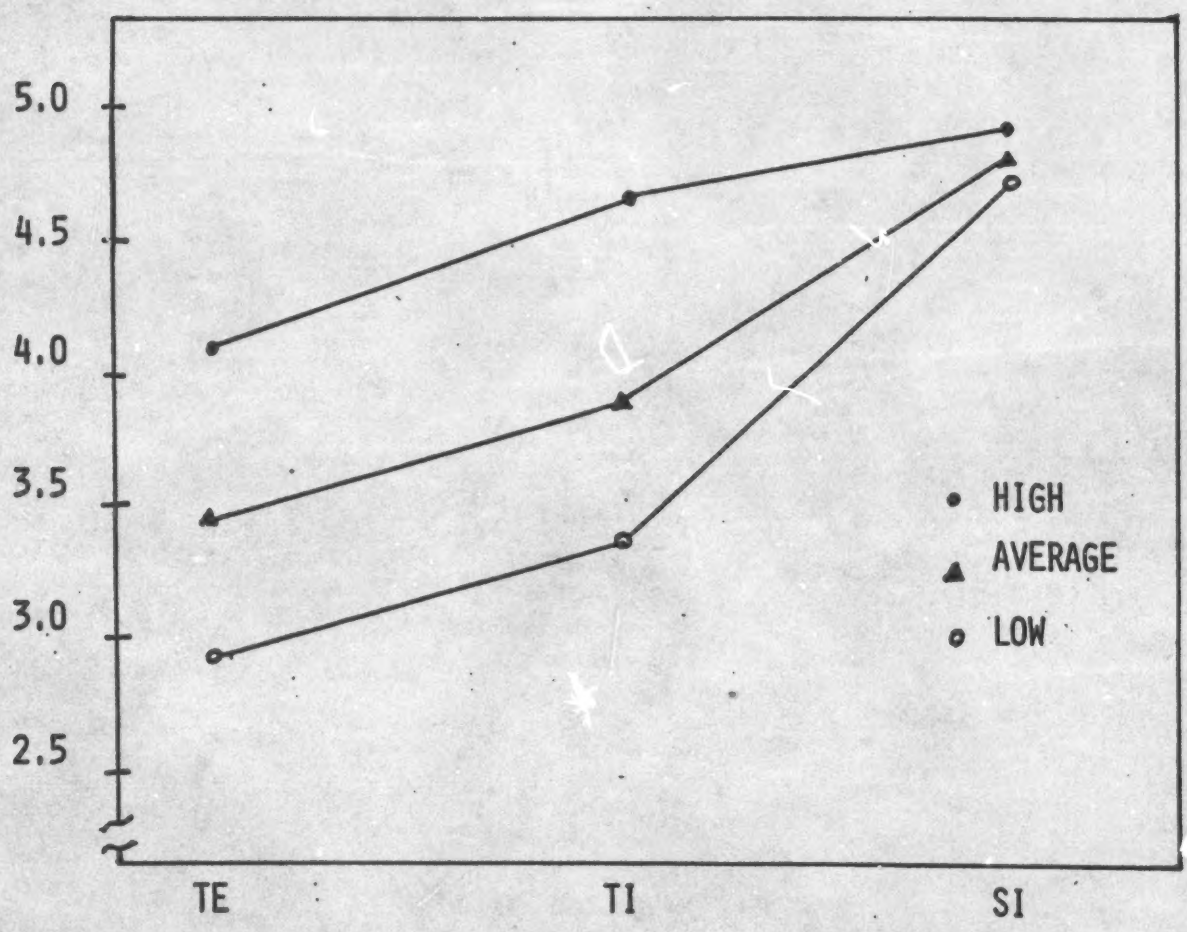

FIGURE 2. TREATMENT X ABILITY INTERACTION, HITS, PASSAGE C.

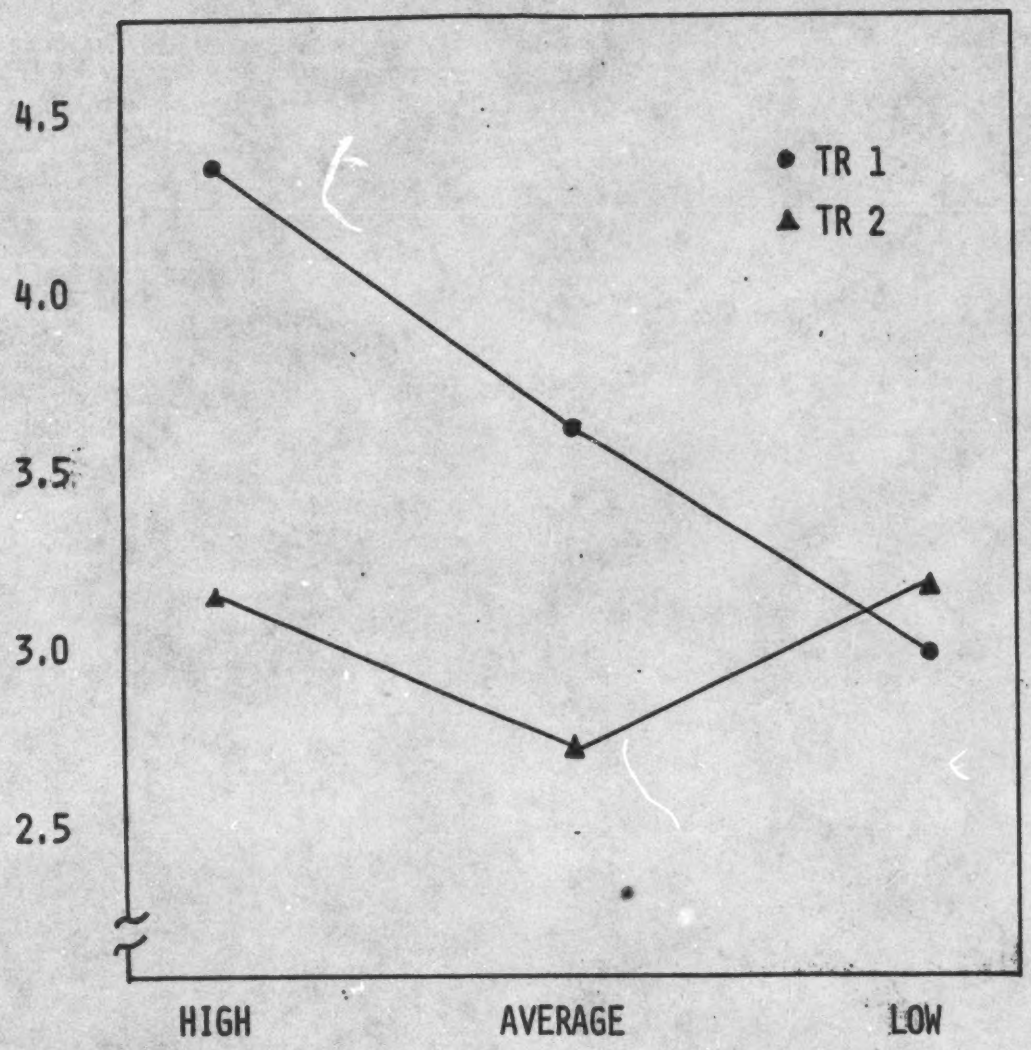

FIGURE 3. TREATMENT X ABILITY INTERACTION, CORRECT RESPONSES, PASSAGE D.